Article

The Use of Artificial Intelligence in the Assessment of User Routes in Shared Mobility Systems in Smart Cities

Andrzej Kubik

Department of Road Transport, Faculty of Transport and Aviation Engineering, Silesian University of Technology, Krasińskiego 8 Street, 40-019 Katowice, Poland; andrzej.kubik@polsl.pl

Abstract: The use of artificial intelligence in solutions used in smart cities is becoming more and more popular. An example of the use of machine learning is the improvement of the management of shared mobility systems in terms of assessing the accuracy of user journeys. Due to the fact that vehicle-sharing systems are appearing in increasing numbers in city centers and outskirts, and the way vehicles are used is not controlled by operators in real mode, there is a need to fill this research gap. The article presents a built machine learning model, which is a supplement to existing research and is updated with new data from the existing system. The developed model is used to determine and assess the accuracy of trips made by users of shared mobility systems. In addition, an application was also created showing an example of using the model in practice. The aim of the article is therefore to indicate the possibility of correct identification of journeys with vehicles from shared mobility systems. Studies have shown that the prediction efficiency of the data generated by the model reached the level of 95% agreement. In addition, the research results indicate that it is possible to automate the process of evaluating journeys made in shared mobility systems. The application of the model in practice will facilitate management and, above all, it is open to further updates. The use of many machine learning models will allow solving many problems that will occur in an increasing number of smart cities.

Keywords: smart cities; shared mobility; machine learning; artificial intelligence; mobility modeling

1. Introduction

In recent years, we have witnessed a significant technological revolution covering not only societies and all areas of the economy, but also areas of public access, such as cities. The digitization of numerous services or products forces contemporary cities to use more and more new IT solutions for their proper functioning [1]. A solution related to this trend is the concept of smart cities, which has been known for more than two decades. This idea assumes the use of various types of information and communication technologies by cities to increase the interactivity and efficiency of the urban infrastructure and its components, as well as to raise the awareness of residents [2].

The emergence of new Smart Cities is a challenge in which designers must take into account areas such as the environment, management, energy consumption, education and human safety [2]. Smart cities are also smart societies that require solutions that will help ensure high quality and adapt to the fast pace of their lives and, at the same time, the ability to carry out personal matters, which can often be located in different urban areas requiring the use of an appropriate and efficient means of transport [3]. Then, one of the domains of smart cities is to provide the society with efficient movement, i.e., the fulfillment of one of their basic needs, which is mobility. In the case of smart cities, this mobility is based on a smart transport network, and the areas of its mobility include aspects such as traffic management, sustainable mobility and also intelligent searching for routes or parking spaces [4]. Among sustainable mobility services, the leading solutions in smart cities and, at the same time, those most accessible to the public, are shared mobility systems [5,6].

These types of services allow sharing journeys or vehicles while providing the opportunity to travel in an alternative way to individual motorists. Shared mobility, in addition to a number of advantages, such as the elimination of transport barriers, is also associated with challenges faced by cities themselves, transport managers, and service providers of individual transport solutions [7]. Among the research works identified in the literature, scientists devote their considerations to the possibilities of solving numerous problems with shared mobility, such as [8–11]:

- adapting shared mobility services to the needs of society;
- selection of vehicles for systems;
- elimination of communication chaos related to vehicles scattered in urban areas;
- modeling the functioning of shared mobility services;
- optimization of shared mobility services;
- security of shared mobility services and others.

Focusing on one of the thematic areas of research on shared mobility, safety, it should be stated that despite the multitude of works related to this issue, the emphasis is on research topics that would concern solving the problem of inappropriate behavior of users using shared mobility system [12]. Numerous scientific studies on the analysis of the functioning of shared mobility systems in cities have noted that users of systems such as scooter-sharing, bike-sharing or car-sharing often use them in a random, often aggressive way, not paying attention to the style of their driving and not caring about rented vehicles [13,14]. As a result, users are often involved in various types of accidents or accidents. Relating this type of behavior to the issue of road safety, research emphasizes that aggressive driving is currently one of the leading problems [15]. Studies show that the aggressive driving style of drivers increases the risk of an accident 15-fold, regardless of whether the driver is a professional or a private driver [16]. Moreover, inappropriate behavior of users of shared mobility directly affects operators providing shared mobility services. They are associated with additional costs incurred for vehicle repairs, but they are also important in perceptions of the services by society at large, which, due to the inappropriate behavior of service users, begins to see shared mobility systems as dangerous [17]. As a result, we are increasingly dealing with some urban centers that limit the functioning of shared mobility systems in their area or create various types of referenda aimed at developing solutions for the proper management of services; such a referendum took place, for example, in Paris [18]. An exemplary speed course of a vehicle used in the carsharing system is shown in Figure 1.

As can be seen in Figure 1, the average speed of a vehicle moving in the city was 31 km/h, which is generally acceptable. Unfortunately, speeding can be also observed in a certain range, where the maximum speed was 81 km/h, which is unacceptable in urban conditions.

To solve the problem of improper use of shared mobility systems, a machine learning model that will support shared mobility operators and smart city managers is proposed by the author to predict the correctness of journeys made by shared mobility vehicles. It is an innovative solution that uses the latest technology, which is artificial intelligence. In addition, the use of artificial intelligence in solutions supporting the functioning of shared mobility systems is a kind of novelty, as there are no such practical applications and theoretical framework in the literature so far. This article is a presentation of the model developed based on research on shared mobility systems carried out by the author. The aim of the work is to indicate the possibility of correct identification of vehicle journeys from shared mobility systems. The need to search for various types of solutions to eliminate the inappropriate behavior of users of shared mobility systems is evidenced by the continuous development of systems that predict that by 2027 the number of users using shared mobility systems will reach five billion [19]. Moreover, according to the report [20], it was estimated that in 2023 over 29 billion devices will be connected to the Internet of Things (IoT). This only underlines the importance of how much data today's cities will have to deal with, especially in terms of the security of using them.
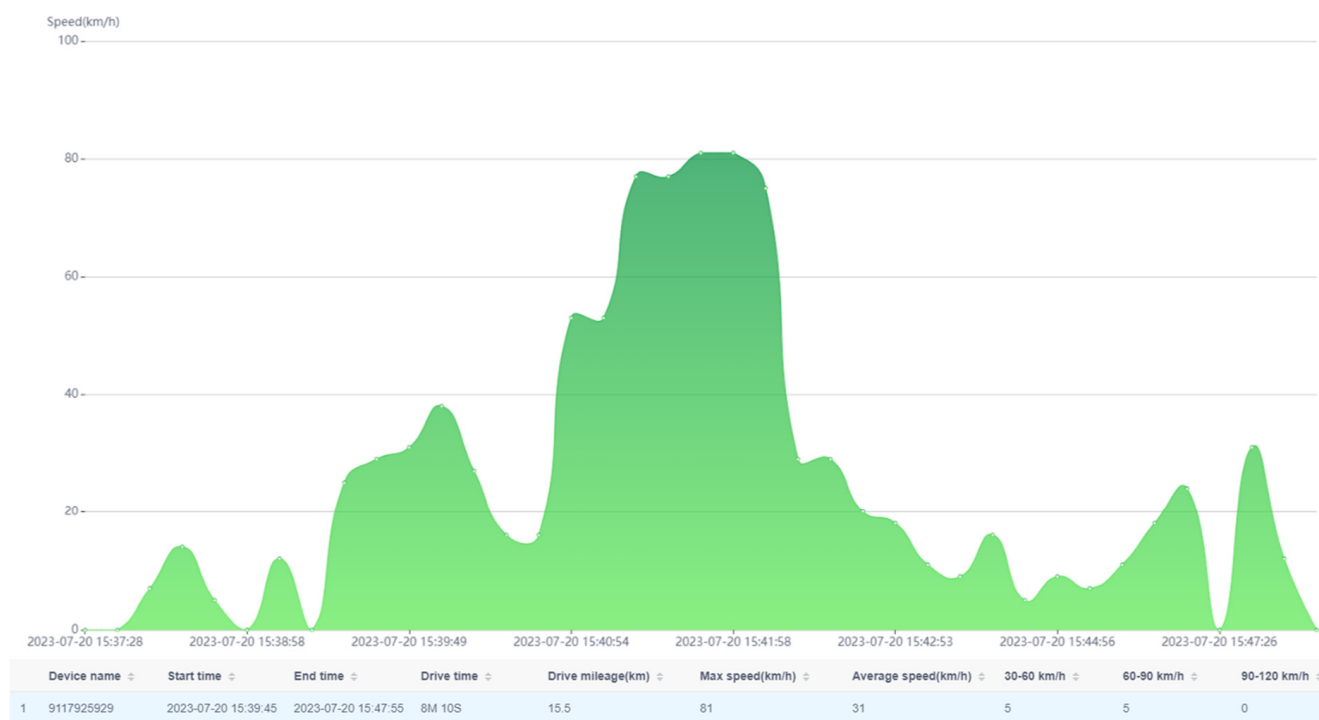
Figure 1. Exemplary speeds of a vehicle moving in urban conditions.

This article has been divided into five parts. The first part is a general outline of the existing problem and a presentation of preliminary information in the field of mobility in smart cities. The second part is a presentation of the methodology that was used to develop the model. The third part presents the results of the obtained research, which are discussed in the fourth part of the work. The fifth part is a summary of the article along with conclusions, limitations and further research plans of the author in the research topic undertaken.

2. Data and Methods

Machine learning is one element of artificial intelligence (AI), which is being used in more and more areas of daily life [21]. The services offered as part of smart cities are also built with the use of artificial intelligence. Previous records of models describing the occurring phenomena required a lot of time to conduct research, and their results were limited by an untested element [22]. One of the solutions currently in use is the ability to use existing data as part of the input data, or data validating the model. In this article, the data used to train the functions published in other original articles [23–27] were also used to verify the data obtained during the test runs.

The proposed model to evaluate the correctness of the driver's journeys was built according to the scheme of its functioning presented in Figure 2.

The developed model was built of four main parts. The first part is the identification of the particular vehicle that has moved in the vehicle-sharing scheme. The data model describing the $M_{vehicle}$ was written with the Formula (1):

$$M_{vehicle}(I_d) = [P_1, P_2, P_3 \ldots, P_n] \tag{1}$$

where:

I_d—individual number of the vehicle [-];
n —number of variables [-];
$P_1, P_2, P_3 \ldots, P_n$—variables describing the technical data of the vehicle [-].
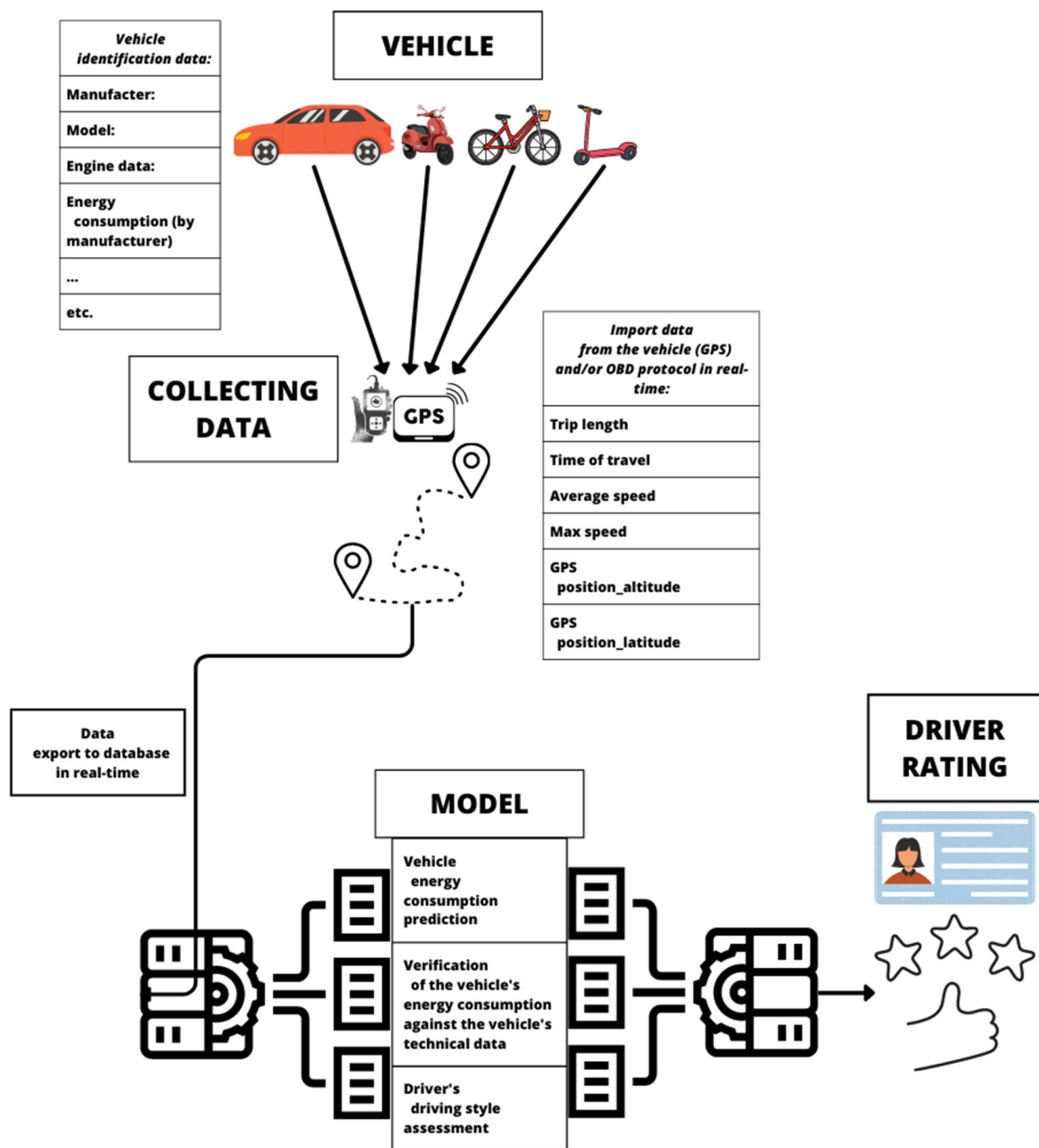
Figure 2. Diagram of the prediction model.

The following exemplary technical data of one of the vehicles (2) were adopted in the developed model:

$$M_{vehicle}(I_d) = [P_1, P_2, P_3, P_4, P_5, P_6] \tag{2}$$

where:

P_1—vehicle brand [-];
P_2—vehicle model [-];
P_3—engine power [kW];
P_4—declared energy consumption [kWh/100 km];
P_5—battery charge status [%];
P_6—number of vehicle seats occupied [-].

Another part of the model is a set of data from each trip (M_{trip}) of the vehicle, represented by a data matrix (3):

$$M_{trip} = \begin{bmatrix} M_{vehicle}(I_d)_{11} & T_{12} & \dots & T_{1j} & I_{user} \\ \vdots & & \vdots & \vdots & \vdots & \vdots \\ M_{vehicle}(I_d)_{i1} & T_{i2} & \dots & T_{ij} & I_{user} \end{bmatrix} \tag{3}$$

where:

i, j —number of horizontal and vertical matrix levels [-];

T_{ij}—data describing the route traveled [-];

I_{user}—individual identification number of a given user [-].

The sample trip data are presented in the form of Table 1.

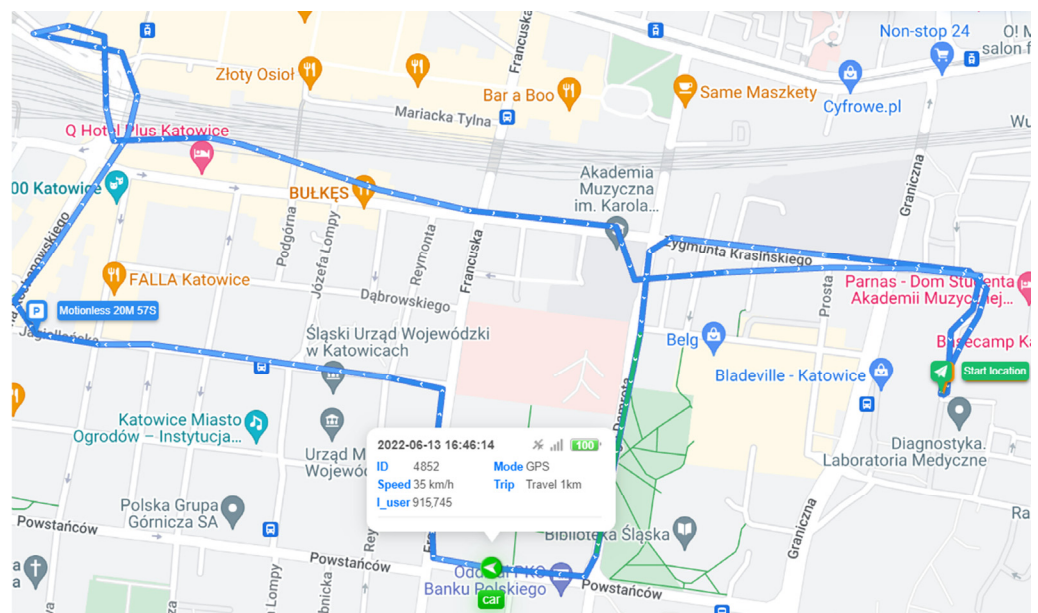An exemplary view of the route taken by the user is shown in Figure 3.

Figure 3. An exemplary graphical interpretation of the vehicle's route.

Subsequently, the $M_{vehicle}(I_d)$ and M_{trip} data are exported to the database, which keeps a copy of the archived data and at the same time analyzes on this basis, identifying the correctness of the trip. The analyzed data are a set of D_{input} training data, which consists of the following elements (4):

$$D_{input} = \{M_{vehicle}, M_{Trip}, M_{rating}\} \tag{4}$$

where:

M_{rating} is an assessment of the drivers' journeys, which was classified by the model (5):

$$M_{rating} = \begin{cases} 1 - vehicle\ driven\ inconsistently \\ 2 - vehicle\ moves\ correctly\ in\ urban\ conditions \\ 3 - vehicle\ moves\ correctly\ in\ extra - urban\ conditions \end{cases} \tag{5}$$

Table 1. Sample route values M_{trip}.

Sample	$M_{vehicle}$ (I_d)	T_{i2}	T_{i3}	T_{i4}	T_{i5}	T_{i6}	T_{i7}	T_{i8}	T_{i9}	T_{i10}	T_{i11}	T_{i12}	T_{i13}	T_{i14}	T_{i15}	T_{i16}	U_{user}
1	5929	18 June 2023 12:24	18 June 2023 12:27	181	1.7	74	33	10	6	4	0	20	19.1873	50.4005	19.2042	50.4121	911,792
2	5414	18 June 2023 18:36	18 June 2023 18:41	242	2.7	64	32	16	16	2	0	34	19.3326	50.5783	19.3041	50.5607	905,840
3	8115	20 June 2023 10:15	20 June 2023 10:18	180	2	100	55	6	4	9	1	20	19.1057	50.0910	19.1309	50.0838	911,745

where: T_{i2}—rental start date [-]; T_{i2}—rental end date [-]; T_{i3}—rental time [s]; T_{i4}—trip length [km]; T_{i5}—Max speed [km/h]; T_{i6}—average speed [km/h]; T_{i7}—number of GPS samples in the speed range of 0–30 km/h [-]; T_{i8}—number of GPS samples in the speed range of 30–60 km/h [-]; T_{i9}—number of GPS samples in the speed range of 60–90 km/h [-]; T_{i10}—number of GPS samples in the speed range of 90–120 km/h [-]; T_{i11}—number of GPS samples in the speed range of >120 km/h [-]; T_{i12}—the number of all GPS samplings during the rental [-]; T_{i13}, T_{i15}—vehicle GPS position data at the start of the rental [-]; T_{i14}, T_{i16}—vehicle GPS position data at the end of the rental [-].

Due to the high availability of model machine learning tools, an analysis of the selection of the optimal model training was carried out for the prediction model [28,29]. Models created on the basis of the methods were adopted for the analysis: Tree models, Naïve Bayes models, Support Vector Machines (SVM) models, K-nearest neighbors (KNN) models, Ensemble (Tree) models and Neutral Network models. In addition, to optimize each of the applied machine learning methods, the minimum value of the classification error was verified for various parameters of model building. Figures 4–9 show the results of individual methods and the values of the minimum classification error in individual iterations (changes in model-building data).
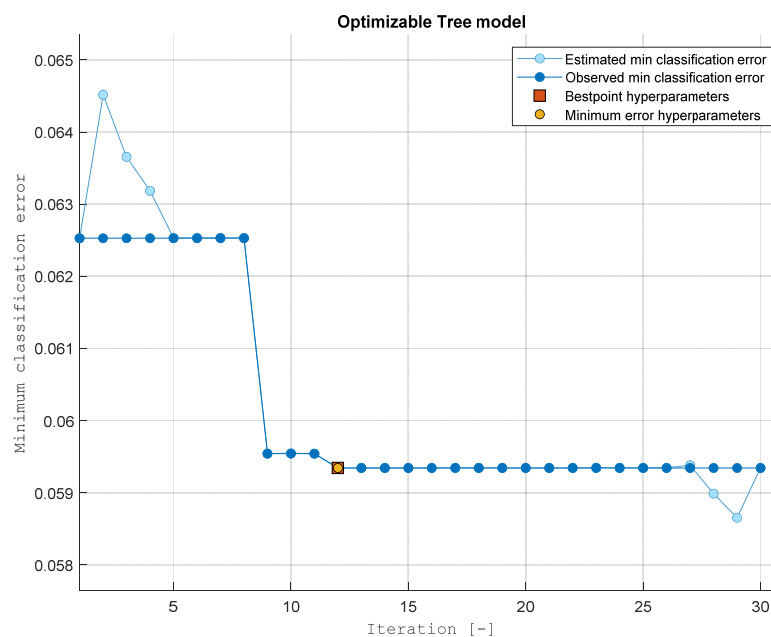

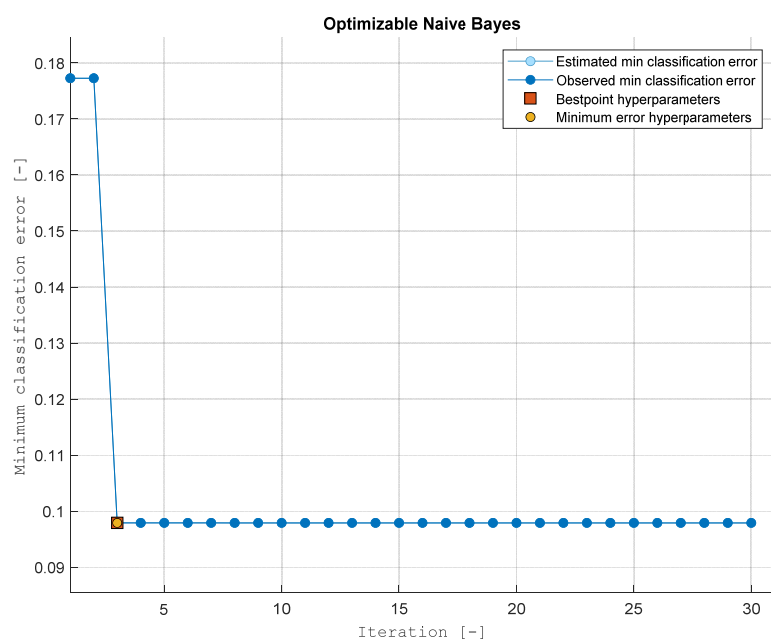**Figure 4.** Changes to the minimum classification error value for the Tree models method.

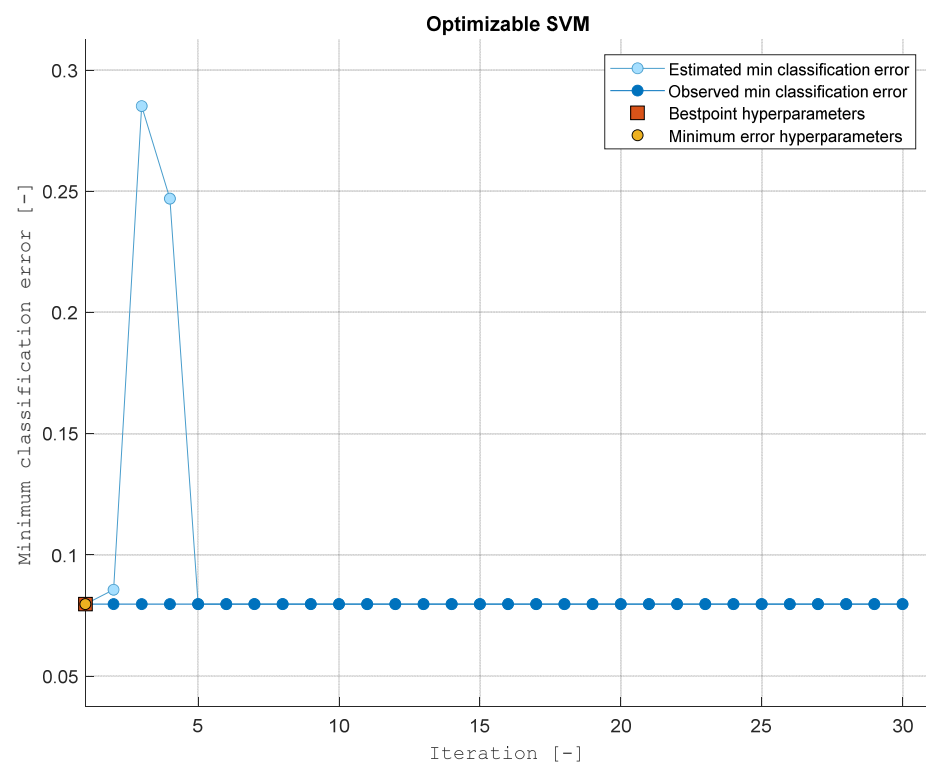**Figure 5.** Changes in the minimum classification error value for the Naïve Bayes models.

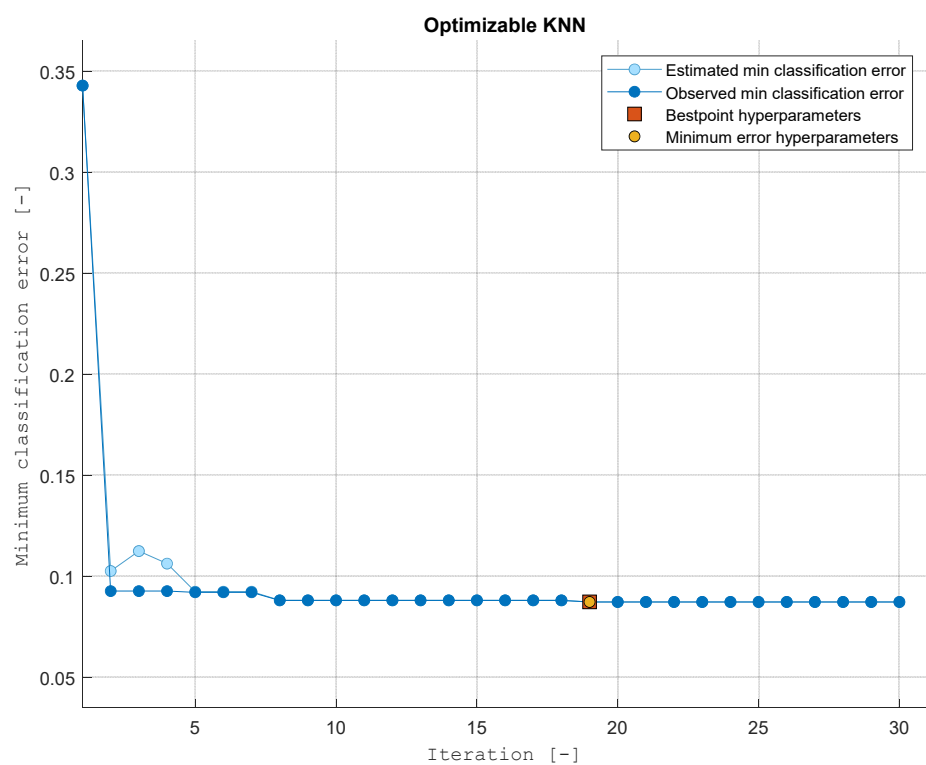**Figure 6.** Changes in the value of the minimum classification error for the BVM method.

Figure 7. Changes in the value of the minimum classification error for the KNN method.
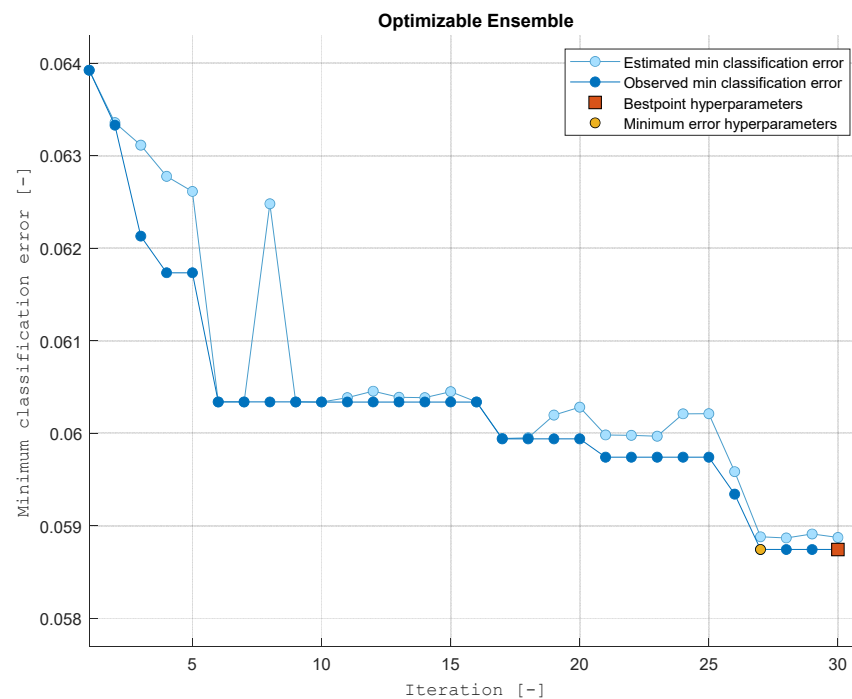
Figure 8. Changes to the minimum classification error value for the Ensemble (Tree) method.

Figure 9. Changes in the value of the minimum classification error for the Neural Network models method.

In order to assess the usefulness of the methods and to create an appropriate model, the assessment measures of individual factors were used, such as: True Positive Rate (TPR), True Negative Rate (TNR), False Positive Rate (FPR), False Negative Rate (FNR), Minimum Accuracy (MA) and minimum value of the classification error (MCE) according to Formulas (6)–(11):

$$TPR = \frac{T_P}{T_P + F_N} \tag{6}$$

where:

T_P—True Positive—the number of observations correctly classified into the positive class.
T_N—True Negative—the number of observations correctly classified into the negative class.
F_N—False Negative—the number of observations classified into the negative class, when in fact the observation belongs to the positive class.

$$TNR = \frac{T_N}{T_N + F_P} \tag{7}$$

where:

F_P—False Positive—the number of cases that are positively classified when in fact the case is negative.

$$FPR = \frac{F_P}{T_N} \tag{8}$$

$$FNR = \frac{F_N}{T_P} \tag{9}$$

$$MA = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{10}$$

$$MCE = 1 - MA \tag{11}$$

After performing the calculations, the lowest value of the change in the minimum value of the classification error for the training data was obtained for the Ensemble (Tree) method, which achieved a result of 5.87%. The second in the ranking was the Tree Models method, which also achieved a result of less than 6% of the classification error. Detailed results for the training data are presented in Table 2. It should be emphasized that the remaining methods after several iterations showed a constant value of the classification error, which ultimately confirmed that the Ensemble (Tree) method is the most advantageous method for further use. The detailed results of the classification error value and the correctness of the classification of the training data model are presented in Table 2.

Table 2. The results of the classification error value and the correctness of the model classification with the training data.

Method	MCE [%]	MA [%]
Tree model	5.94	94.06
Naïve Bayes	9.79	90.21
SVM	8.32	91.68
KNN	8.72	91.28
Ensemble (Tree)	5.87	94.13
Neural Network	10.01	89.99

The models created were trained with training data to verify the correctness of the models' predictions. For the models, after checking with test data, the values of the minimum classification error of the models and their classification accuracy were recalculated. Detailed results of the classification error value and the correctness of the model classification for the test data are presented in Table 3.

Table 3. The results of the classification error value and the correctness of the model classification for the test data.

Method	MCE [%]	Model Accuracy [%]
Tree model	6.60	93.40
Naïve Bayes	9.20	90.80
SVM	7.90	92.10
KNN	8.40	91.60
Ensemble (Tree)	6.30	93.70
Neural Network	9.70	90.30

It can be concluded that after testing the model with test data, the accuracy of the model prediction decreased by about 0.5%. Due to the obtained high rates of prediction accuracy and low value of the classification error, the Ensemble (Tree) method was chosen as the main research method.

The selected method also has several advantages and disadvantages. The main advantage is that the proposed method is characterized by better accuracy and efficiency in the case of complex solutions. In addition, the method can work with different types of data, and it is not required to enter data of one type. An undoubted disadvantage of the method is the need for high computing power so that the study is not time-consuming. The work is carried out on a large amount of data that is necessary for the model to work. With a large number of variables, it may be difficult to interpret the results and determine the correct relationships. However, despite the disadvantages, in the analyzed example this method obtained the highest value of matching the training data to the test data [30,31].

The next step is to perform a graphical interpretation of the efficiency of the predictive model by determining the qualitative characteristics of the classifiers. This graphical interpretation is called the ROC curve. The ROC curve is created by determining the cut-off points (k_{od}) according to the Formula (12):

$$k_{od} \rightarrow (x(k_{od}), \ y(k_{od})) = (FPR(k_{od}), TPR(k_{od})) \tag{12}$$

Then, partial dependence graphs will be determined, showing the impact of the individual factors (features) tested on the predicted result of the prediction model. The partial dependence function can be represented by the Formula (13):

$$\hat{f}_s(x_s) = E_{Xc}\left[\hat{f}_s(x_s, \ X_C)\right] = \int \hat{f}(x_s, X_C)dP(X_C) \tag{13}$$

where:

x_s—features for which the dependency function is determined;
X_C—other features used in the prediction model—random variables.

The end result of the prediction model will be the determination of the correctness of the journey made and the identification of whether the vehicle is moving in urban or extra-urban conditions. The developed machine learning model will use a group of supervised machine learning factors based on the ensemble tree method.

3. Results

In order to build the final version of the model, data from the existing vehicle sharing system operating in Poland were used. For this purpose, 5558 individual journeys of users were analyzed and individual journeys were assessed. The D_{input} input data were divided into two groups of data: model training data (90%—of the D_{input} input data set) and model training data—Input data (10%—of the D_{input} input data set).

Based on the D_{input} data, a confusion matrix was determined to determine the TPR, FRN, PPV and FDR values. The confusion matrix is shown in Figures 10 and 11.

Confusion matrix - Optimizable Ensemble

True Class [-]	Predicted: 1	Predicted: 2	Predicted: 3	TPR	FNR
1	94.7%	2.1%	3.2%	94.7%	5.3%
2	6.7%	86.7%	6.7%	86.7%	13.3%
3	1.4%	0.9%	97.7%	97.7%	2.3%

Predicted Class [-]

Figure 10. Confusion matrix with true positive rates and false negative rates.

Confusion matrix - Optimizable Ensemble

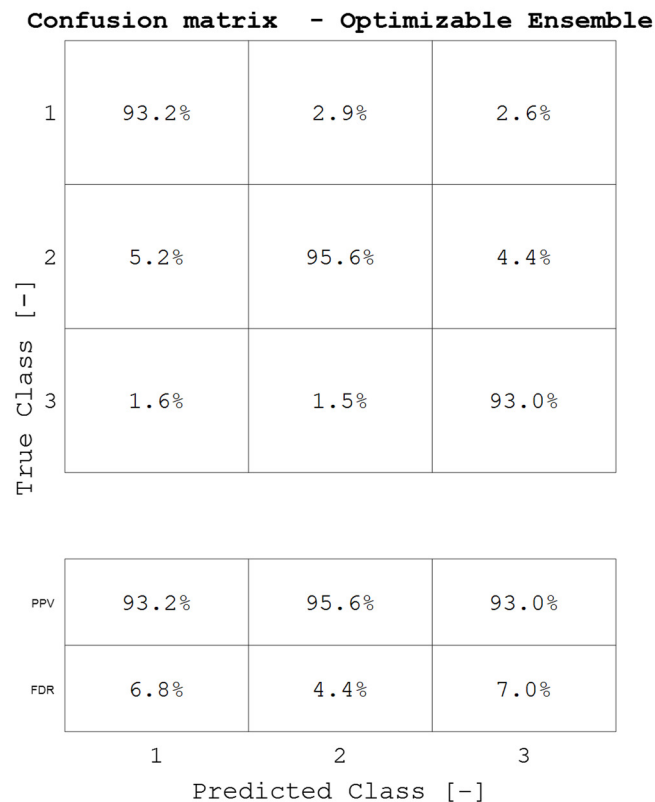| True Class [-]	Predicted: 1	Predicted: 2	Predicted: 3
1	93.2%	2.9%	2.6%
2	5.2%	95.6%	4.4%
3	1.6%	1.5%	93.0%
PPV	93.2%	95.6%	93.0%
FDR	6.8%	4.4%	7.0%

Predicted Class [-]

Figure 11. Confusion matrix with positive predictive values and false discovery rates.

As can be seen, the coefficient values obtained in Figures 10 and 11 are characterized by high values above 90%. This means that the training data and test data have been correctly classified. Subsequently, a graphical interpretation of the efficiency was determined by creating the ROC curve. The ROC curve for the created model is shown in Figure 12.
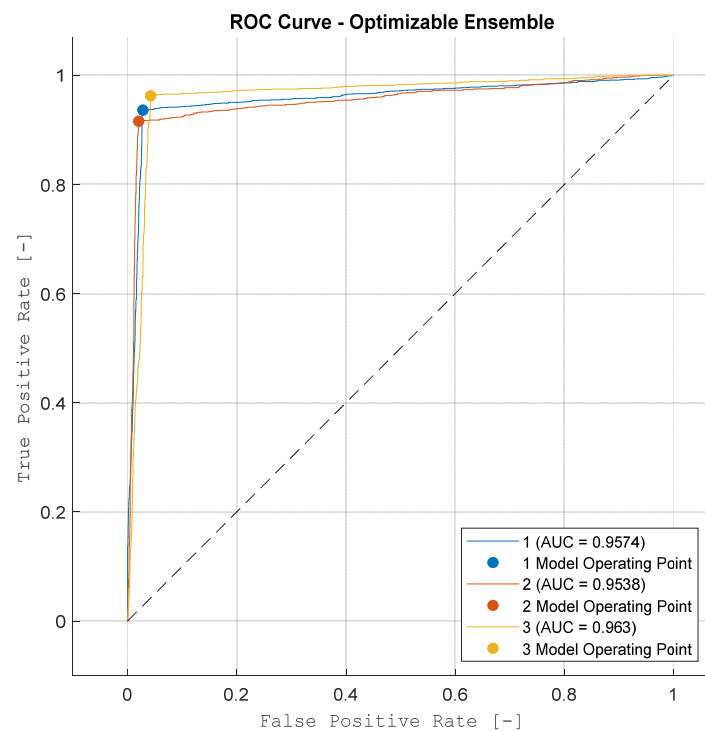
Figure 12. ROC Curve for prediction model.

As can be seen in Figure 12, the variability of TPR depending on FPR is linear. In addition, the Gini coefficient (AUC) for all three predictors is >95%, which only underlines the correctness of the prediction model.

Then, in Figures 13–18, graphs of partial dependence of the classifiers forming the model are presented—Markings on the graphs:

- red line—prediction of vehicle movement in urban conditions;
- yellow line—prediction of vehicle movement in non-urban conditions;
- blue line—incorrect prediction of vehicle movement.

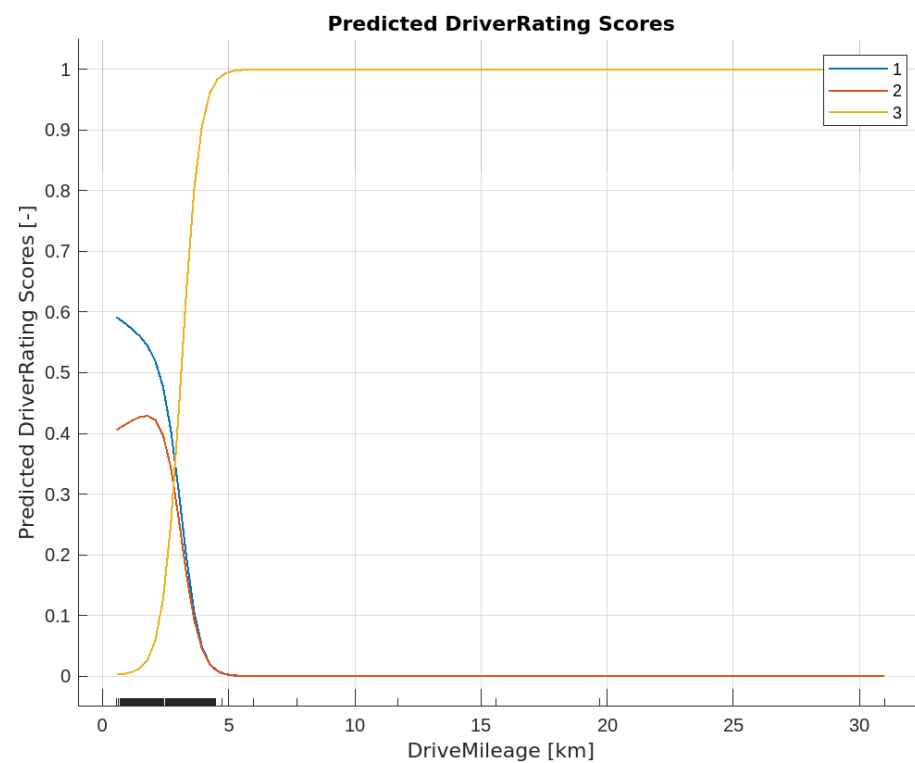**Figure 13.** The dependence of the prediction model on the DriveTime variability function.

Figure 14. The dependence of the prediction model on the DriveMileage variability function.
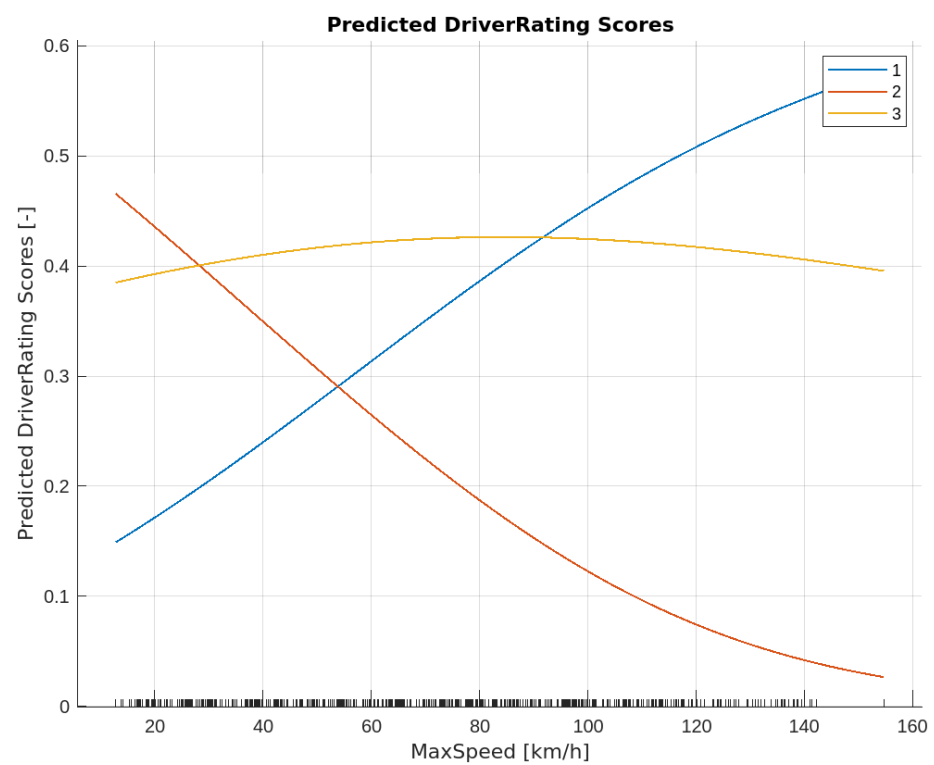

Figure 15. The dependence of the prediction model on the MaxSpeed variability function.
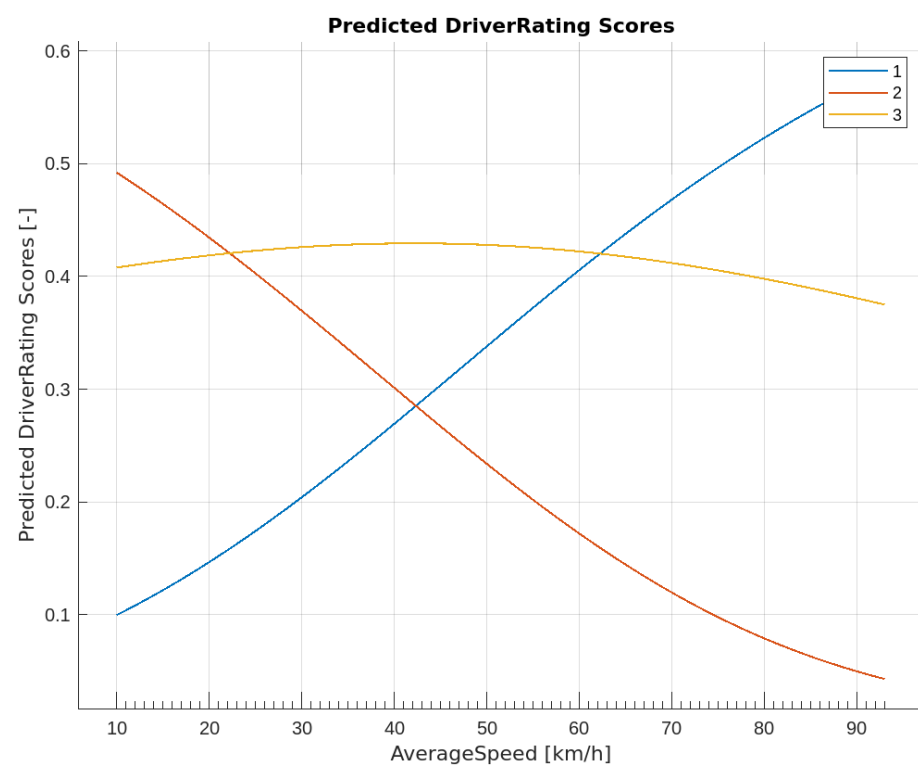
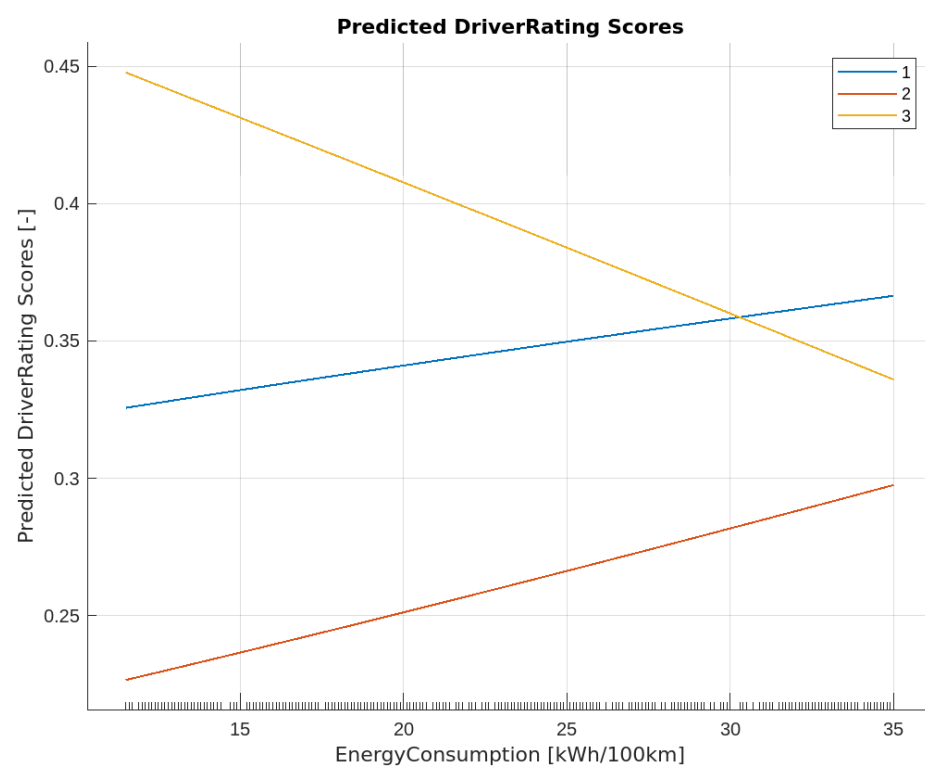**Figure 16.** Dependence of the prediction model on the AverageSpeed variability function.

Figure 17. Dependence of the prediction model as a function of EnergyConsumption variability.
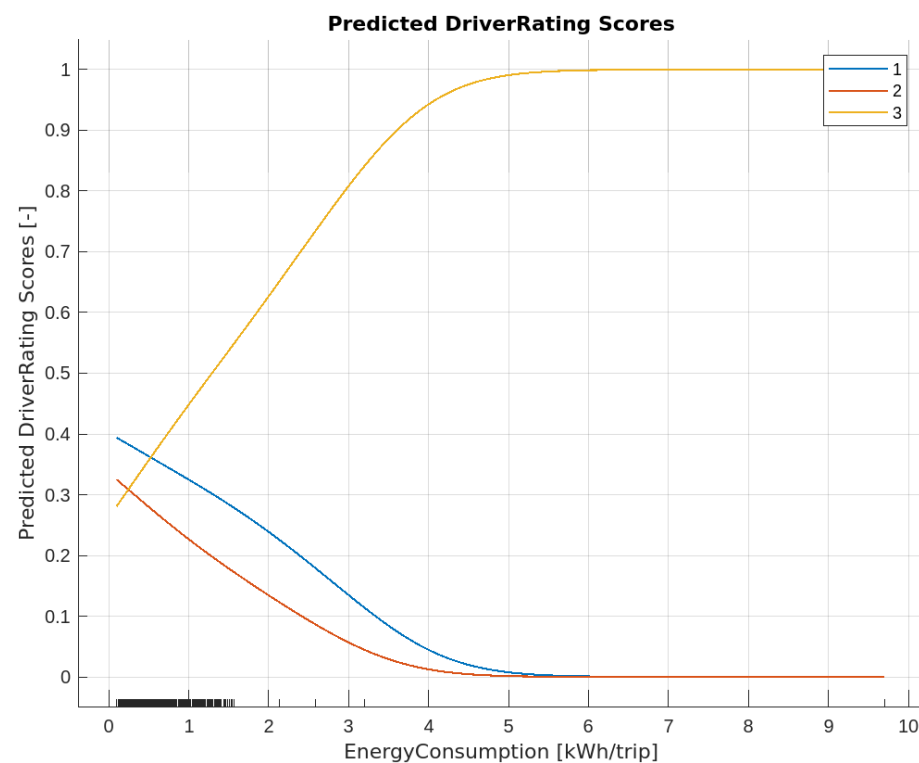
Predicted DriverRating Scores

Figure 18. Dependence of the prediction model as a function of Energy Consumption/trip variability.

The indication of the vehicle travel time (Figure 13) allows one to determine that, as the travel time increases, the probability of discrepancies detection decreases. On the other hand, the probability of moving outside the built-up area increases. The probability of moving around the city assumes a practically constant value of 26%.

The relationship between the length of the trip (Figure 14) and the prediction is very important. In the range up to 3 km, the model determines that vehicles mainly move in the urban cycle, correctly and incorrectly. However, after exceeding the value of 3 km, the prediction of movement in the non-urban area begins to dominate.

The maximum speed detection dependence (Figure 15) has a very large impact on the prediction result. As can be seen in Figure 15, if the speed exceeds 54 km/h, the model will strongly strive to predict incorrect vehicle movement in urban conditions. The stability of the prediction is maintained by detecting movement outside the built-up area.

Determining the average speed (Figure 16) with which the vehicle moved during the entire route determines that when the speed exceeds 45 km/h, the model prediction changes from correct vehicle movement to incorrect.

The graph of the partial influence of the factor 'average value of energy consumption' (Figure 17) on the prediction value shows that with the increase in the average value of consumption, the value of the average energy consumption decreases in the case of moving in non-urban conditions. On the contrary, the average energy consumption increases in the case of incorrect trips.

The last factor analyzed was energy consumption in relation to the route traveled—shown in Figure 18. The factor constructed in this way determines the nature of the factor, and it is visible that with the increase in this factor, the nature of driving the vehicle in extra-urban mode is predicted. On the other hand, low values of the factor characterize incorrect driving of the vehicle.

Based on the built model, an application with an intuitive graphical interface was launched, which has an implemented prediction model for evaluating individual driver trips. An example screenshot of the application is shown in Figure 19. The application has two basic functioning dependencies:

- input data can be imported directly from the vehicle—in real mode;
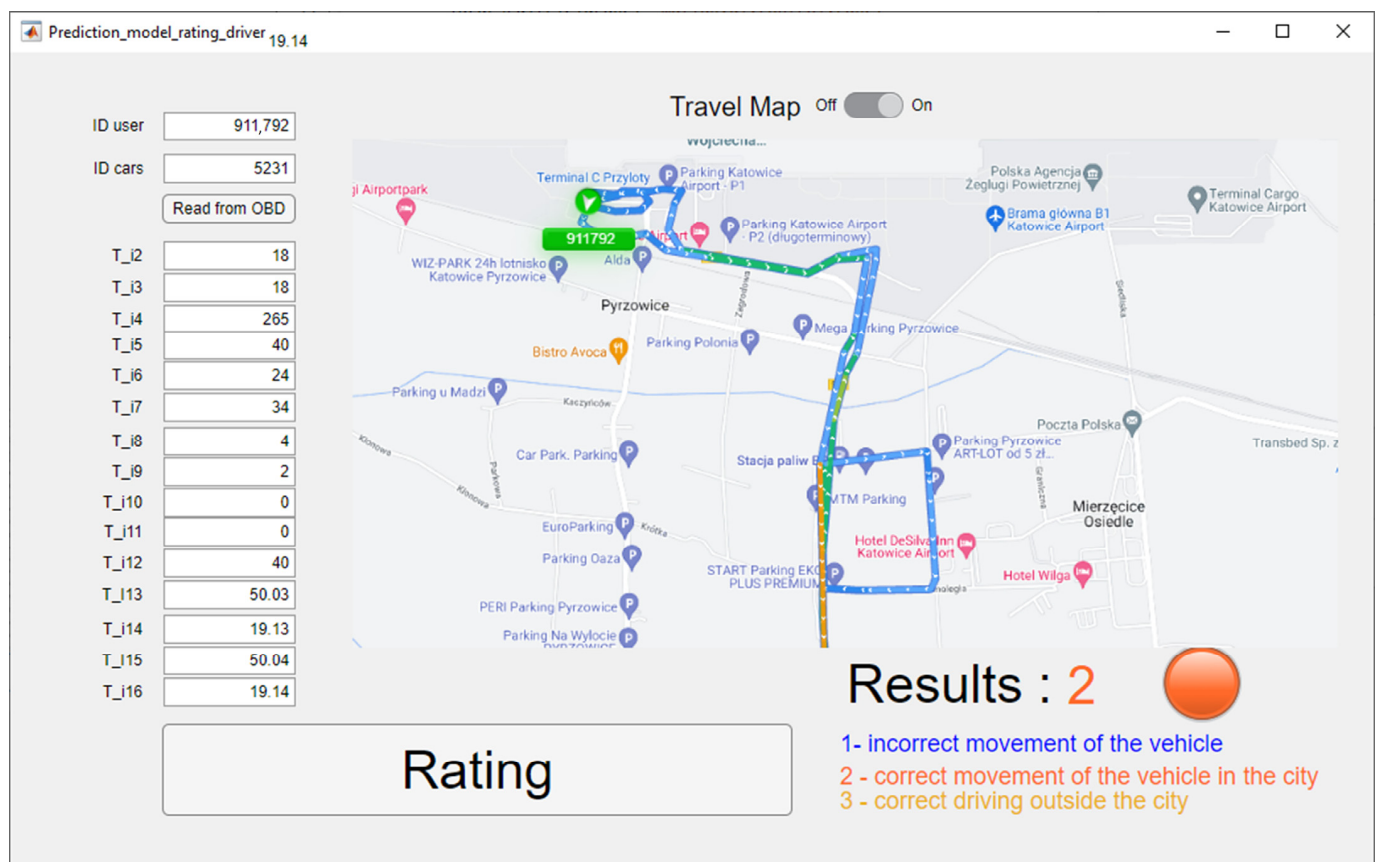- archival data can be kept—manually.

Figure 19. The use of an application to evaluate the journeys of drivers in a mobility sharing system.

In the application window, it is possible to enter data or their parts in accordance with Table 1. The model does not require the completeness of all data to predict the result. However, it should be remembered that entering more data increases the chances of obtaining the correct result.

The final effect of the application is to determine whether, based on the received data, the driver used the vehicle correctly or whether he or she misused the vehicle. The application was made using the Matlab2023a software.

4. Discussion

The results obtained under the developed model clearly indicate that it is possible to build a fully functional tool to predict the behavior of vehicle drivers in shared mobility systems. The input data and functions of the behaviors defined in previous studies enabled the creation of a model that generates correct results of classification of incorrect behaviors of drivers at the level of 95%. The conducted research is an innovative solution that allows determining the behavior of drivers of vehicles from short-term rentals. In addition, the model also received a graphical overlay that allows people who are not necessarily well-versed in the subject of machine learning to use the software.

When comparing the research conducted with the analyses performed by other scientists, it should be stated that no identical studies referring directly to shared mobility have been reported. Referring to other research carried out in the field of broadly understood transport, it is worth mentioning the research carried out by Mohammadnazar et al. [32], which concerned the development of a proposal for an unsupervised machine learning model as one of the elements of identifying groups of drivers and the variability

of driving behavior. However, it was a division that distinguished only drivers, highway drivers, street drivers and local road drivers. The model developed by the authors [30] did not take into account driving identification of whether the vehicle was moving in accordance with the regulations. In addition, it used another method of unsupervised machine learning, k-means.

For comparison, Liu et al. [33] in their research attempted to use machine learning to predict the driving conditions of an electric and a hybrid vehicle. As a result, they proved that the simulations carried out using the proposed model can provide an improvement in speed prediction, which in turn will translate into improved vehicle operating economy. This is extremely important in the case of vehicles used in shared mobility systems, whose operational activities are very often neglected.

Subsequently, Wang et al. [34] used machine learning to predict passenger injuries. They proved that it is possible to identify a collision scenario with a delay of 1.2 ms. This type of application will improve the traffic safety of autonomous vehicles. They used the SVM algorithm to build it.

The safety of driving a vehicle in urban conditions is a function of a large number of factors and variables. Each vehicle driver has his own individual driving style, which can be applied or can be an act of temporary preference of the driver [32]. Deng et al. [34] proposed the Markov model as a method of recognizing the driving style, which uses data on the braking characteristics of the vehicle. The research consisted of the subjective assessment of the drivers who underwent the experiment; while making the journey, they interpreted what driving style they used. Unfortunately, the identification of one's own passage is burdened with a significant error. Aggressive driving is one of the road safety problems. Studies [35] indicate that the aggressive driving style of drivers increases the risk of an accident 15-fold, regardless of whether the driver is a professional or private driver [15,16].

The integration of many prediction models of phenomena that occur in shared mobility services may contribute to the creation of a global model that will treat existing models as subordinate. Tuncel et al. [36] attempted to integrate ride-matching and vehicle-balancing models for shared mobility services. The integrated model solves the problem of matching passages and balancing vehicles. Such a connection leads to an improvement in the functioning of the entire system and also allows the detection of vehicles that are not often used in the shared system.

Due to the fact that machine learning models process large amounts of data, it is possible to use data training of a given model for other applications. The authors [37] determined that estimating safety using pre-programmed dependencies can increase the traffic safety of unprotected road users, which is extremely important in the structures of smart cities.

5. Conclusions

In conclusion, the variety of emerging artificial solutions will support the management and functioning of modern cities. The resulting machine learning model, which can assess the journey of a particular user in real time and on the basis of archive data, is a support for shared mobility operators. Unfortunately, at the moment, crossings are not controlled and more and more often users use vehicles contrary to the regulations and local regulations.

The created model generates the evaluation values of the correctness of the run and classifies which run was made with an accuracy of over 95%, which makes it a very accurate and stable model. In addition, operators can pay attention to which factors generate the highest probability of a given classification.

Based on the research conducted and the created prediction model for the assessment of the correctness of the ride in shared mobility systems, the following conclusions were formulated:

- On the basis of the obtained partial graphs of dependencies of the prediction model, it is possible to determine the probability of the influence of the factor on the evaluation of the trip;
- The practical use of the model together with the graphic overlay facilitates the use of the built prediction model;
- The obtained accuracy of the model prediction is at a very high level—over 95% of the correctness of the classification of passages;
- Machine learning models can be expanded; however, it is important not to overtrain the model with incorrect data;
- The use of machine learning in security-related applications should be performed using supervised machine learning.

The built model can be updated with data from other previous studies or studies by other authors. Therefore, it is possible, based on data on journeys and their classification, to update individual journey prices or to block the possibility of using the system by per-sons who excessively do not comply with the regulations or who break local regulations.

The conducted research on shared mobility services points to current and future directions that the author wants to take.

There is a huge research gap in the use of machine learning in solutions that can automate the work of shared mobility systems. In addition, machine learning will reduce the need to analyze huge amounts of data to detect inconsistencies. From the point of view of the practitioner—the operator and the policy—of local authorities, it will simplify the management and functioning of the system. Furthermore, the model can become a subsystem of existing solutions used in smart cities.

In the next stage, the model will be tested in practice on a test number of vehicles from the fleet of the functioning sharing system. The application graphic page will also be updated, where the user service model will be implemented, along with their travel history.

The limitation of the model's application, at the moment, is the possibility of using it only in the mode without the possibility of data training. Model training data update the model only in manual mode overseen by the model author. Another limitation is also the need to have computing equipment that will analyze the data received on an ongoing basis. In addition, the model should support, and not exclude, the human factor in decision making at the current stage.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Herath, H.M.K.K.M.B.; Mittal, M. Adoption of Artificial Intelligence in Smart Cities: A Comprehensive Review. *Int. J. Inf. Manag. Data Insights* **2022**, *2*, 100076. [CrossRef]
2. Ingwersen, P.; Serrano-López, A.E. Smart City Research 1990–2016. *Scientometrics* **2018**, *117*, 1205–1236. [CrossRef]
3. Arfat, Y.; Aqib, M.; Mehmood, R.; Albeshri, A.; Katib, I.; Albogami, N.; Alzahrani, A. Enabling Smarter Societies through Mobile Big Data Fogs and Clouds. *Procedia Comput. Sci.* **2017**, *109*, 1128–1133. [CrossRef]
4. Bellini, P.; Nesi, P.; Pantaleo, G. IoT-Enabled Smart Cities: A Review of Concepts, Frameworks and Key Technologies. *Appl. Sci.* **2022**, *12*, 1607. [CrossRef]
5. Czech, P.; Turoń, K.; Urbańczyk, R. Bike-Sharing as an Element of Integrated Urban Transport System. In *Advanced Solutions of Transport Systems for Growing Mobility*; Sierpiński, G., Ed.; Advances in Intelligent Systems and Computing; Springer International Publishing: Cham, Switzerland, 2018; Volume 631, pp. 103–111. [CrossRef]

6. Golba, D.; Turoń, K.; Czech, P. Diversity as an Opportunity and Challenge of Modern Organizations in TSL Area. *SJSUTST* **2016**, *90*, 63–69. [CrossRef]

7. Nansubuga, B.; Kowalkowski, C. Carsharing: A Systematic Literature Review and Research Agenda. *JOSM* **2021**, *32*, 55–91. [CrossRef]

8. Alonso-Almeida, M.D.M. To Use or Not Use Car Sharing Mobility in the Ongoing COVID-19 Pandemic? Identifying Sharing Mobility Behaviour in Times of Crisis. *Int. J. Environ. Res. Public Health* **2022**, *19*, 3127. [CrossRef] [PubMed]

9. Guyader, H.; Friman, M.; Olsson, L.E. Shared Mobility: Evolving Practices for Sustainability. *Sustainability* **2021**, *13*, 12148. [CrossRef]

10. Nawangpalupi, C.; Demirbilek, O. Investigation on the Drivers and the Barriers for Travel Behaviour Changes and the Analysis of the Impact: A Case Study of Car Sharing in Australia. *Int. J. Environ. Cult. Econ. Soc. Sustain. Annu. Rev.* **2008**, *4*, 1–10. [CrossRef]

11. Carrese, S.; D'Andreagiovanni, F.; Giacchetti, T.; Nardin, A.; Zamberlan, L. A Beautiful Fleet: Optimal Repositioning in E-Scooter Sharing Systems for Urban Decorum. *Transp. Res. Procedia* **2021**, *52*, 581–588. [CrossRef]

12. Mayo, F.L.; Maglasang, R.S.; Moridpour, S.; Taboada, E.B. Impact of Transport Policies to Commuter Safety in Urban Cities of a Developing Country: A Sustainability and System Perspective. *Case Stud. Transp. Policy* **2022**, *10*, 2138–2152. [CrossRef]

13. Turoń, K.; Kubik, A.; Chen, F.; Wang, H.; Łazarz, B. A Holistic Approach to Electric Shared Mobility Systems Development—Modelling and Optimization Aspects. *Energies* **2020**, *13*, 5810. [CrossRef]

14. Turoń, K.; Kubik, A.; Chen, F. What Car for Car-Sharing? Conventional, Electric, Hybrid or Hydrogen Fleet? Analysis of the Vehicle Selection Criteria for Car-Sharing Systems. *Energies* **2022**, *15*, 4344. [CrossRef]

15. Dingus, T.A.; Guo, F.; Lee, S.; Antin, J.F.; Perez, M.; Buchanan-King, M.; Hankey, J. Driver Crash Risk Factors and Prevalence Evaluation Using Naturalistic Driving Data. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 2636–2641. [CrossRef] [PubMed]

16. Stephens, A.N.; Newnam, S.; Young, K.L. Who's More Angry? Comparing Driver Anger and Aggression within Work and Personal Driving Contexts. *Transp. Res. Part F Traffic Psychol. Behav.* **2021**, *81*, 408–416. [CrossRef]

17. Matyja, T.; Kubik, A.; Stanik, Z. Possibility to Use Professional Bicycle Computers for the Scientific Evaluation of Electric Bikes: Trajectory, Distance, and Slope Data. *Energies* **2022**, *15*, 758. [CrossRef]

18. McFarland, M. Scooters Are a Huge Problem for Cities. No One Knows How to Solve It Yet. Available online: https://edition.cnn.com/2019/08/30/tech/scooter-management/index.html (accessed on 30 May 2023).

19. Shared Mobility—Worldwide. Available online: https://www.statista.com/outlook/mmo/shared-mobility/worldwide (accessed on 30 May 2023).

20. European Commission; Joint Research Centre. *AI Watch, AI Uptake in Smart Mobility*; European Commission: Luxembourg, 2021.

21. Ejaz, W.; Anpalagan, A. *Internet of Things for Smart Cities: Technologies, Big Data and Security*; Springer: Berlin/Heidelberg, Germany, 2019.

22. Soo, A.; Wang, L.; Wang, C.; Shon, H.K. MachIne Learning for Nutrient Recovery in the Smart City Circular Economy—A Review. *Process Saf. Environ. Prot.* **2023**, *173*, 529–557. [CrossRef]

23. Kubik, A. CO_2 Emissions of Electric Scooters Used in Shared Mobility Systems. *Energies* **2022**, *15*, 8188. [CrossRef]

24. Kubik, A.; Turoń, K.; Folęga, P.; Chen, F. CO_2 Emissions—Evidence from Internal Combustion and Electric Engine Vehicles from Car-Sharing Systems. *Energies* **2023**, *16*, 2185. [CrossRef]

25. Kubik, A. Impact of the Use of Electric Scooters from Shared Mobility Systems on the Users. *Smart Cities* **2022**, *5*, 1079–1091. [CrossRef]

26. Turoń, K.; Kubik, A.; Chen, F. Operational Aspects of Electric Vehicles from Car-Sharing Systems. *Energies* **2019**, *12*, 4614. [CrossRef]

27. Kubik, A. The Energy Consumption of Electric Scooters Used in the Polish Shared Mobility Market. *Energies* **2022**, *15*, 8193. [CrossRef]

28. Omar, A.; Delnaz, A.; Nik-Bakht, M. Comparative Analysis of Machine Learning Techniques for Predicting Water Main Failures in the City of Kitchener. *J. Infrastruct. Intell. Resil.* **2023**, *2*, 100044. [CrossRef]

29. Chen, X.; Shuai, C.; Zhao, B.; Zhang, Y.; Li, K. Imputing Environmental Impact Missing Data of the Industrial Sector for Chinese Cities: A Machine Learning Approach. *Environ. Impact Assess. Rev.* **2023**, *100*, 107050. [CrossRef]

30. Kelleher, J.D.; Mac Namee, B.; D'Arcy, A. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*, 2nd ed.; The MIT Press: Cambridge, MA, USA, 2020.

31. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; Adaptive computation and machine learning; The MIT Press: Cambridge, MA, USA, 2016.

32. Mohammadnazar, A.; Arvin, R.; Khattak, A.J. Classifying Travelers' Driving Style Using Basic Safety Messages Generated by Connected Vehicles: Application of Unsupervised Machine Learning. *Transp. Res. Part C Emerg. Technol.* **2021**, *122*, 102917. [CrossRef]

33. Liu, Y.; Li, J.; Gao, J.; Lei, Z.; Zhang, Y.; Chen, Z. Prediction of Vehicle Driving Conditions with Incorporation of Stochastic Forecasting and Machine Learning and a Case Study in Energy Management of Plug-in Hybrid Electric Vehicles. *Mech. Syst. Signal Process.* **2021**, *158*, 107765. [CrossRef]

34. Wang, Q.; Gan, S.; Chen, W.; Li, Q.; Nie, B. A Data-Driven, Kinematic Feature-Based, near Real-Time Algorithm for Injury Severity Prediction of Vehicle Occupants. *Accid. Anal. Prev.* **2021**, *156*, 106149. [CrossRef]

35. Deng, C.; Wu, C.; Lyu, N.; Huang, Z. Driving Style Recognition Method Using Braking Characteristics Based on Hidden Markov Model. *PLoS ONE* **2017**, *12*, e0182419. [CrossRef]
36. Tuncel, K.; Koutsopoulos, H.N.; Ma, Z. An Integrated Ride-Matching and Vehicle-Rebalancing Model for Shared Mobility on-Demand Services. *Comput. Oper. Res.* **2023**, *159*, 106317. [CrossRef]
37. Hula, A.; De Zwart, R.; Mons, C.; Weijermars, W.; Boghani, H.; Thomas, P. Using Reaction Times and Accident Statistics for Safety Impact Prediction of Automated Vehicles on Road Safety of Vulnerable Road Users. *Saf. Sci.* **2023**, *162*, 106091. [CrossRef]